|  |  |
| --- | --- |
| **Practicum Case** |  |
| SCIE6062 | SCIE6062001 | SCIE6062016 | SCIE6062049  Computational Biology |
| **Mathematics & Statistics** | **E231-SCIE6062-VO01-01** |
| ***Valid on*** *Even Semester Year 2022/2023* | **Revision 00** |

**Learning Outcome**

* LO3 – implement basic bioinformatics analysis in scope of DNA composition and sequence analysis, protein synthesis, sequence alignment, and other works related to biological database using Biopython
* LO4 – analyze the basic bioinformatics analysis results using Biopython

## Topic

* Session 01 – Sequence Analysis using BioPython

## Sub Topics

* BioPython: Introduction and Installation
* Working with Sequence
* Sequence and Subsequence Manipulation
* Nucleotide Frequency Plot

## Soal

*Case*

**Sequence Analysis using BioPython**

For given **DNA sequences**:

Sequence A:

**AGCTTGCAGCGTCCGTTAGCTCGAGTCCAGGACGTTAGTCCTGCAGTC**

Sequence B:

**CAGTAAGTTGCCGTTAGCGCGTAGTGCCAGTAAGCGGCTCGTTAGTGG**

Please use **Biopython** to:

* Find the **length** of **both sequences**.
* Determine the **number of times** the **codon** (**CGC**) appears in **both sequences**.
* Locate the **first occurrence** **index** **position** of the **sub-sequence** (**CAGTC**) in **both sequences**.
* **Combine** the **first 15 nucleotides** from **sequence A** and the **last 10 nucleotides** from **sequence B** to **create** a new DNA sequence called **sequence C**.
* **Reverse** the sequence of **sequence C**.
* **Plot** the **nucleotide base** (**A, C, G, T**) **frequency** of each of the **sequence A**, **sequence B**, and **sequence** **C** using the **Matplotlib** library.